

日本国特許庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日  
Date of Application: 2003年 4月18日

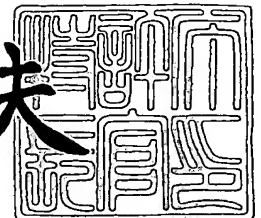
出願番号  
Application Number: 特願2003-113750  
[ST. 10/C]: [JP2003-113750]

出願人  
Applicant(s): 中山 光雄

2004年 5月14日

特許庁長官  
Commissioner,  
Japan Patent Office

今井康夫



出証番号 出証特2004-3040899

【書類名】 特許願

【整理番号】 D2526-NAK

【提出日】 平成15年 4月18日

【あて先】 特許庁長官 殿

【国際特許分類】 G06K 9/22

【発明者】

    【住所又は居所】 東京都港区芝浦 4 - 1 3 - 3    トリニティ芝浦 1 8 0 6

    【氏名】 中山 光雄

【特許出願人】

    【住所又は居所】 東京都港区芝浦 4 - 1 3 - 3    トリニティ芝浦 1 8 0 6

    【氏名又は名称】 中山 光雄

【代理人】

    【識別番号】 100086368

    【弁理士】

    【氏名又は名称】 萩原 誠

【手数料の表示】

    【予納台帳番号】 041793

    【納付金額】 21,000円

【提出物件の目録】

    【物件名】 明細書    1

    【物件名】 図面    1

    【物件名】 要約書    1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 光学端末装置およびシステム

【特許請求の範囲】

【請求項 1】 書類上の意図した情報領域を光学的にスキャニングしてイメージ情報として取り込むイメージキャプチャー部と、

取込んだイメージ情報およびスキャン中に検出したイメージの位置情報とを処理する画像処理部と、

前記イメージキャプチャー部によりスキャニングしたイメージデータを外部に出力する出力部と、

前記出力部により外部に出力されたイメージ技術の文字認識結果を入力する入力部と、

前期入力部で入力した文字認識結果を表示する表示部とを有し、

前記画像処理部はスキャン中に取込んだ複数のイメージとその各々のイメージの中から選ばれた小イメージから位置座標情報を計算することを特徴とする光学端末装置。

【請求項 2】 前記光学端末装置は前記書類上の文字情報の中から文字認識を行う領域を任意に指定できるように前記書類上を移動可能であり、

さらにパーソナルコンピュータから入力する入力部を備えており、

前記イメージデータは前記出力部より前記パーソナルコンピュータの入力部に出力され、前記イメージデータが前記パーソナルコンピュータにて前記位置情報データに基づいて全体のイメージデータが形成され、そのイメージを文字認識しその結果を前記入力部から取込むことを特徴とする請求項 1 に記載の光学端末装置。

【請求項 3】 前記光学端末装置は、ポインティングデバイス機能を持つフリーハンドスキャナーであることを特徴とする請求項 2 に記載の光学端末装置。

【請求項 4】 前記画像処理部は、イメージ上の対向する位置に少なくとも 2ヶ所設けられた小イメージの領域から位置座標情報を計算することを特徴とする請求項 3 に記載の光学端末装置。

【請求項 5】 前記光学端末装置は、高速双方向通信バスにより前記パーソ

ナルコンピュータと接続されることを特徴とする請求項1、2、3、4に記載の光学端末装置。

【請求項6】 前記光学端末装置は前記表示部に表示された文字認識結果を確定する確定手段を有し、前記指定した領域の文字認識結果は前記表示部にリアルタイムに表示され、前記表示部にリアルタイムに表示された文字認識の開始箇所と終了箇所の位置を前記確定手段で確定するまでは修正することを可能とする請求項2に記載の光学端末装置。

【請求項7】 前記光学端末装置は、光あるいは電波等の無線手段によって前記パーソナルコンピュータと接続されることを特徴とする請求項1、2に記載の光学端末装置。

【請求項8】 前記光学端末装置は携帯電話の一部として組み込まれ携帯電話の無線公衆回線を介してサーバーと接続され、前記サーバに前記光学端末装置にて取得した前記イメージデータを送信することを特徴とする請求項1に記載の光学端末装置。

【請求項9】 請求項8に記載の光学端末装置を備えた携帯電話。

【請求項10】 前記携帯電話は、無線あるいは高速双方向通信バスを介してパーソナルコンピュータに接続され、前記携帯電話の前記光学端末装置をポインティングデバイスとして使用することを特徴とする請求項9に記載の携帯電話。

【請求項11】 書類上の意図する領域を光学端末装置を用いてスキャンし取込んだイメージ情報を画像処理して合成する方法であって、

(1) イメージキャプチャ部により書類上の任意に指定した領域をスキャンし複数のイメージを取込むステップと、

(2) 前記取込んだ複数のイメージの各々のイメージデータの対角の位置に少なくとも2ヶ所の位置座標計算用の領域を指定しメモリに取り込むステップと、

(3) 上記位置座標データから相対的な位置座標を計算するステップと、

(4) 上記イメージデータを位置座標データとともに出力部から送信するステップと、

(5) 上記位置座標データに基づいてイメージ情報をつなぎ合わせ、全体のイメ

ージデータを形成するステップと  
を具備したことを特徴とする画像処理合成方法。

【請求項 1 2】 携帯電話の一部として組み込まれた請求項 8 に記載の光学  
端末装置から携帯電話の無線公衆回線を介して前記サーバに送られる前記光学端  
末装置にて取得した前記イメージデータをサーバー内のソフトにて処理すること  
を特徴とする通信システム。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】

本発明はイメージスキャナ等の画像入力手段を持つ端末および携帯端末、また  
これを用いた光学的文字認識装置およびそのシステムに関するものである。

より具体的には、書類上の任意の領域の文字を端末あるいは携帯端末によって  
光学的イメージとしておれなく確実に取込みを行い、その取込んだイメージに対  
して文字認識を行ってパーソナルコンピュータのアプリケーションソフトに文字  
コードデータとしてデータ入力する、あるいは遠隔地にある装置で文字認識を行  
ってデータを装置の所望の制御部に入力する、またそのデータを利用することを  
可能とした光学的文字認識装置およびそのシステムに関する。

【 0 0 0 2 】

【従来の技術】

従来のパーソナルコンピュータとフラット・ベッドあるいはシート・フィード  
・スキャナ（以下スキャナと称す）とを用いた印刷書類（以下、単に書類と称す  
）の文字認識装置では、印刷文字の認識作業において以下のような操作を必要と  
する。

（操作 1）・・・書類スキャン 先ず、スキャナへ書類をセットし、書類全体を  
スキャンする。

（操作 2）・・・文字認識スキャナから送られてきた書類全体のイメージは、パ  
ーソナルコンピュータのディスプレイ上に展開され、文字認識ソフトで表示され  
る。文字認識ソフト上で“文字認識させたい範囲”（以下、“意図する領域”と  
称す）をマウスで指定してから、文字認識ソフトを操作して、“意図する領域”

のイメージデータをテキストデータ（文字コードデータ）に変換し、再度ディスプレイ上に表示する。

（操作 3）・・・修正 文字認識ソフトの認識結果は、文字認識ソフトに表示された書類のイメージと認識結果であるテキストデータとを突き合わせて、キーボードを用いて修正をする。

（操作 4）・・・アプリケーションへのコピーまたはペーストを行い、次に、この修正されたテキストデータを、マウスを用いてコピーし、ワープロや表計算機能をもつアプリケーションソフトにペーストする。

（反復操作）・・・認識させたい書類が複数ある場合は、前記操作 1 から操作 4 を繰り返す。書類内に“意図する領域”が複数箇所ある場合は、操作 2 から操作 4 までを繰り返す。

#### 【 0 0 0 3 】

フラット・ベッドあるいはシート・フィード・スキャナではガイドレールがありスキャン時にぶれを生じないような対策が取られている。フラット・ベッドあるいはシート・フィード・スキャナ以外には、必要な部分のイメージだけを取込み文字認識するハンディタイプのスキャナがある。このようなハンデイスキャナではスキャン部の受光部面積が小さいため文字認識を実施したい領域の全部をスキャンできない場合が多く、スキャナー自体を左右に動かして“意図する領域”全体の取込みを行う必要がある。

#### 【 0 0 0 4 】

しかし、ハンデイスキャナを使用する場合には、手ぶれによりイメージがうまく修得されず文字認識がされないという問題点があった。

そこで、上記問題点を解決すべく、特許文献 1 には書類上の“意図する領域”をイメージスキャナでスキャンすることにより得られたイメージデータを、パーソナルコンピュータ内の文字認識ソフトによりテキストデータに変換し、アプリケーションソフトに直接入力すること、また、“意図する領域”の入力開始位置の指定と確認はイメージスキャナの手元の LCD により行うことを可能にするハンデイスキャナおよびこれを用いた光学的文字認識装置が開示されている。

#### 【 0 0 0 5 】

**【特許文献 1】**

国際公開 WO 0 0 / 2 6 8 5 1

**【0 0 0 6】****【発明が解決しようとする課題】**

しかしながら、上記従来技術のようにガイドレールが無いハンディスキャナの場合、書類上の“意図する領域”をスキャンする際にいわゆる手ぶれが生じ、画像の歪みの原因となっていた。

**【0 0 0 7】****【発明の目的】**

したがって、本発明は、所望の領域をスキャンする際のいわゆる手ぶれによる画像の歪みを解消し、完全なイメージデータを取得することによって、文字認識を確実に実行できるようにするとともに、取込んだイメージデータに対する文字認識を端末あるいは携帯端末以外の装置において実行し、その認識結果をアプリケーションあるいは遠隔地にある装置で文字認識を行ってデータを装置の所望の制御部に入力することを可能とした光学的文字認識装置およびシステムを提供することを目的とする。

**【0 0 0 8】****【課題を解決するための手段】**

そこで、上記課題を解決するため、本発明によるフリーハンドスキャナは、書類上の文字情報の中から文字認識を行う領域を任意に指定できるように前記書類上を移動可能であり、書類上の文字情報を光学的にスキャニングして静止イメージ情報として取り込むイメージセンサと、スキャン中に取込んだ複数の静止イメージとその各々の位置座標を計算する画像処理回路と、前記イメージセンサによりスキャニングした静止イメージ情報を外部に出力する出力部と、外部での認識結果を入力する入力部と認識結果を表示する表示部とを有する。

**【0 0 0 9】****【発明の実施の形態】**

以下、添付図面を参照しながら、本発明の実施形態について説明する。図 1 は、本発明による光学的文字認識装置の実施の画像処理の流れ図である。図 2 は、

光学的文字認識装置のシステム構成図である。図3は、イメージセンサ部の撮影方法を説明するための図である。図4は、イメージセンサ部の撮影タイミング図である。図5は、本イメージセンサを用いたマウススキャナのシステム構成図である。図6は、マウススキャナの構造図である。図7は、マウススキャナを組み込んだ携帯電話の構造図である。図8は、スキャナカメラ付携帯電話が役立つ各種インターネットサービスを説明するための図である。

#### 【0010】

なお、本実施の形態の説明中の用語「書類」は、本、雑誌、新聞またはパンフレット、名刺等の文字や文章が記述された印刷原稿を意味する。また、用語「文字認識」は、「かな」、「片仮名」、「漢字」、「英数字」等の一般的な文字の認識を意味するだけではなく、例えばバーコード等のような符号であっても、コンピュータで認識してキャラクタデータに変換可能なイメージデータであれば範疇に含むものとする。

#### 【0011】

(1) 本発明の実施の形態である画像処理フロー

図1を参照しながら、本発明の実施の形態であるフリーハンドスキャナによるスキニングの際のイメージデータ処理を説明する。

まず、本フリーハンドスキャナによって、書類上の情報のスキャンを開始する(S101)。イメージセンサと高速電子シャッターによって静止イメージが連続して取得され(S103)、画像処理手段に送られる。

#### 【0012】

画像処理手段では各々のイメージの位置座標が計算処理され、各イメージは一時保存される(S105)。各イメージは、一定量貯まると、位置座標データとともに、スキャンが終了するまで出力部からパーソナルコンピュータへ送信される(S107~109)。各静止イメージデータは、その位置座標データに基づいて各イメージをつなぎ合わされ、全体のイメージデータが形成される(S111)。

その後、画像のみ利用の場合は、ユーザ利用の任意のアプリケーション上のカーソル位置へ選択した画像データを自動的に(または直接)入力し(S119)



、作業終了となる（S121）。

#### 【0013】

文字を利用する場合は、認識ソフトによる全体静止画像データの文字データへの自動変換を行い（S113）、認識した文字データを一時的に保管する（S115）。つづいて、認識結果（文字情報）の手元表示と関心のある文字の選択確定を行い（S117）、ユーザ利用の任意のアプリケーション上のカーソル位置へ選択した文字データを自動的に（または直接）入力し（S119）、作業終了となる（S121）。

以上によりいわゆる手ぶれが防止でき完全なイメージデータが取得できる。

#### 【0014】

（2）本発明の実施の形態であるフリーハントスキャナのハードウェア構成

図2を参照しながら、本発明の実施の形態であるフリーハントスキャナ（イメージスキャナ）200のハードウェア構成を説明する。

イメージスキャナ200は、書類上の文字情報を光学的にスキャニングして静止イメージ情報として取り込むためのストロボ発光用LED201、およびイメージセンサ203、ストロボライト発生器205、A/Dコンバータ207、マイクロ・コントローラ209、画像合成計算部211、表示部213、USBインタフェース215、メモリ（デジタル画像データ記憶部）217とを具備している。

#### 【0015】

USBインタフェース215は、イメージセンサ203によりスキャニングしたイメージ情報をパーソナルコンピュータ220に出力する出力部と、この出力部によりパーソナルコンピュータ220に出力されたイメージ情報の文字認識結果をパーソナルコンピュータ220から再入力する入力部とを有している。

表示部213は該入力部で入力した文字認識結果を表示し、フリーハントスキャナ200は表示部213で表示した文字認識結果の確定の有無を行う操作部（図示せず）を有する。

#### 【0016】

イメージスキャナ200は、書類上の文字情報の中から文字認識を行う領域を

任意に指定できるように書類上を移動可能である。また、指定した領域の文字認識結果を表示部に表示することによりスキニングした書類上の文字情報の変換結果を手元で確認できるとともに、図示しない操作部により、変換結果の確定が行われるとその旨をパーソナルコンピュータ（パソコン 220）に通知する。

パーソナルコンピュータ 200 は、少なくとも 1 つのアプリケーションソフトウェア 223 と、イメージスキャナ 200 の出力部より出力されたイメージ情報の文字認識を行う文字認識ソフトウェアとを搭載する。そして、イメージセンサ 200 によりスキニングしたイメージ情報を入力すると、文字認識ソフトウェアによる文字認識結果をイメージスキャナ 200 に送信する。

イメージスキャナ 200 より変換結果の確定の通知を入力すると、アプリケーションソフトウェアのポインティングデバイスのカーソルで指定された位置に確定された文字データが入力される。

#### 【0017】

##### （3）本発明の実施の形態であるデータ入力方法

本発明によるデータ入力方法、すなわちパーソナルコンピュータで動作するワープロソフトや表計算ソフトまたはデータベースソフト等のアプリケーションソフトにおけるデータ入力方法の処理シーケンスを以下に説明する。

書類上の文字情報の中から文字認識を行う領域を任意に指定できるように書類上を移動可能なイメージスキャナ 200 によりスキニングすると、スキニングしたイメージデータがパーソナルコンピュータ 220 に送信される。

パーソナルコンピュータ 220 がこのイメージデータを受信すると、データ入力を行うアプリケーションソフトの背後で動作している文字認識ソフトにより、イメージデータの文字認識を行って該当する文字コードデータをイメージスキャナ 200 に送信する。イメージスキャナ 200 が文字コードデータを受信すると、この文字を表示部 213 に表示し、この表示された文字コードデータの確定操作が行われると、パーソナルコンピュータ 220 に文字コードデータの確定を通知する。

パーソナルコンピュータ 220 が文字コードデータの確定の通知を受信すると、アプリケーションソフト 223 にこの文字コードデータが入力される。

## 【0018】

## (4) 本発明の実施の形態であるの撮影方法

図3を参照しながら、本発明の実施の形態であるフリーハントスキャナ200のイメージセンサ部203の撮影方法について説明する。ここでは、画素数が30万画素（570x570画素）のランダムアクセス可能なCMOSイメージセンサを用いて説明するがこれに限られるものではない。

現在CMOSイメージセンサ203は、毎秒30フレームの撮影が可能であり、これは33msec 間隔で撮影することができる。フリーハントスキャナ200では、スキャン速度が毎秒30cmの場合、このイメージセンサ203により1mm間隔で大静止画が撮影される。ここで用いるイメージセンサ203の撮影サイズを10mm×10mmとし、このサイズを大静止画として処理する。また、小静止画のサイズを2mm×2mmとする。

## 【0019】

大静止画の位置座標を決定するための小静止画は、図3に示すように、大静止画の対角の位置の2か所が選ばれる。すなわち大静止画（30万画素（570 x 570画素））内の対角線上の左下のA部分近傍の位置に矩形領域部分2mm角（114×114画素）と、またその対角にあたる右上のB部分近傍の位置に矩形領域部分2mm角（114×114画素）が選ばれる。

スキャン作業中は常に移動量を認識するため、またスキャン作業中に起きる回転角度を算出するため、上記小静止画は常に対として処理される。このとき小静止画の総画素数の50%（6,500画素）から30%の重なり合いが必要である。

## 【0020】

スキャン・ボタンを押してスキャンが開始されると大静止画データが取得され、そのなかの小静止画を用いて求められた大静止画の位置座標（重心位置座標値（X,Y）、回転角度 $\theta$ ）とともに大静止画はパーソナルコンピュータ（PC）220に送られ、PC220では送られてきたこの位置座標データを用いて転送された各静止画をメモリ上に配置してつなぎ合わせる処理が実施される。文字認識処理を行う場合にはPC220にインストールされたアプリケーションソフト223を用いて文字認識処理が実施される。

## 【0021】

文字やコード認識を実施する場合については、対象とする文字やコードに合わせた最適解像度を用いるのが望ましい。かな漢字の認識には1mm当り16画素、英数字では1mm当り8画素の解像度が適していると言われている。

各画素へのランダムが可能なCMOSイメージセンサ203では、ユーザの指定した任意のエリアのみの読み出しを行うことで、読出しフレーム速度の高速化が可能である。このエリアは一般にWOI (Window Of Interest) と言われている。

上記の30万画素(570x570画素)のCMOSイメージセンサ203の解像度は1mm当り57画素に相当するので、文字やコード認識に適した解像度まで、このWOIを用いて下げることにより、センサ203のフレーム速度を上げ、最大スキャン速度を上げることが可能になる。

## 【0022】

撮影時間中に対象とする画像(文字やコード)の解像度以上に、画像が動くと正確な静止画が得られない。1mm当り16画素のかな漢字認識の場合、1ドットのサイズは、0.0625mmであり、撮影中に許される各ドットの移動量は約0.1ドットと考えると、許される撮影時間は、 $0.00625\text{mm} \div \text{スキャン速度 (mm/sec)}$ となる。一例として、スキャン速度が毎秒100mmの場合の撮影時間は、この式より62.5マイクロ秒となる。

## 【0023】

(5) 本発明の実施の形態であるタイミングチャート

図4は、本発明の実施の形態であるイメージセンサ部203の撮影タイミングを模式的に表したものである。スキャナ回路は図2に、全体システムの画像処理の流れは図1に示したものを使用する。大静止画、小静止画は用途別で決められた撮影間隔でデータを取込むものとする。

## 【0024】

(6) 具体的実施例の説明

(6-1) 実施例1

まず、文字やコード認識を実施する場合について説明する。

かな漢字処理用では、1mm当り57画素のセンサにおいて使用する画素を間引

き、解像度を下げて、1mm当り16画素とすると、サイズ2mm角の各小静止画の間隔が1mm（50%重なり）の場合、WOIを利用すると、毎秒90フレーム（撮影間隔11msec）に上げて、最大スキャン速度を毎秒90mmに上げることができる。

各小静止画における照合対象の総画素数は648画素（ $36 \times 36$ （画素） $\times 0.5$ ）と大幅に減少するのでパターン照合のための処理時間も短くなる。次に英数字用では、1mm当り8画素とし、使用する画素を間引き解像度を下げることにより、毎秒150フレーム（撮影間隔6.7msec）に上げ、最大スキャン速度を毎秒150mmとすることができる。各小静止画における照合画素の総数は128画素（ $16 \times 16$ （画素） $\times 0.5$ ）と減少し、さらに短時間でのパターン照合処理が可能となる。

#### 【0025】

次に文字認識の際の動作手順を説明するが、全イメージを取る場合とは処理する画素数が少ないだけで基本的には同じである。

スキャン・ボタンを押し、スキャン開始する。

1-1. まずLED光源201のパルス発光により、大静止画データ全体が撮影されCMOSセンサー203の画素に電荷蓄積される。電荷蓄積に要する時間は対象画面の明度にもよるが60μsecとする。

#### 【0026】

1-2. 次に撮影された大静止画の各画素の信号は、A/D変換器207によりデジタルデータに変換される。文字認識の場合、最低2値データでもよいが、良好な認識率を得るには多値化をするほうが良い。2か所の小静止画の画像データについてはアドレスがスキャンされ、画素データはDSP（Digital Signal Processor）メモリ211bへ転送される。DSPメモリ211bの読み取り速度を60nsecとするとデータ取込みは0.3msec（ $= 60\text{ns} \times 8\text{bit} \times 36 \times 36\text{画素} \times 0.5$ ）で完了する。

#### 【0027】

1-3. DSPメモリ211bへ転送された小静止画データはDSP211によって処理され、大静止画のA、B点の座標値への換算に用いられる。この座標値計

算はフレーム撮影後、次のフレーム撮影までの間（33ms）に行われる。座標値（回転角度）の計算は、前後のフレームの小静止画のパターン照合により行われる。パターン照合は、相関係数を求めると計算時間がかかるので、画素数が少ない場合は短い計算時間ですむ残差法（着目画素とそれを取囲む画素のパターンの差を求める）により行う。

#### 【0028】

1-4. DSP211により計算された大静止画の座標位置（回転角度）と、大静止画の画素データは、メモリ217に一端格納される。この後引き続き次のフレーム撮影が行われる。

1-5. その後、大静止画データが一定量まとまれば、USB（Universal Serial Bus）を用いてPC220へ連続的に転送される。その際、各静止画には位置座標（重心位置座標値、回転角度）データが添付される。

1-6. 上記処理が繰り返された後、スキャン・ボタンを離すことにより、スキャンは終了する。

以上の動作後に、PC220における処理が開始される。

#### 【0029】

2-1. 画像データ合成部227は、スキャナ200から転送された各静止画をその位置座標情報に基づいて、重なり合っている部分を除きながらつなぎ合わせ、メモリ（図示せず）内に記録する。但し、スキャナ200側にメモリおよびDSP計算余力がある場合、静止画のつなぎ合わせ作業はスキャナ200側で行う。

2-2. その後、BMPデータ変換部229がPC用画像フォーマット（BMP）への変換を行う。

以上の処理の後に、文字認識処理が実施される。文字認識ソフトは、ユーザが利用するPC用アプリケーションの背後で動作する形に作られている。

#### 【0030】

3-1. 繋ぎ合わせが終了した静止画は、文字認識ソフト（文字／コード認識部231）に転送される。

3-2. 文字／コード認識部231での文字認識結果は、PC200内の一時

保管メモリ（テキストファイル記憶部233）に送られる。一時保管メモリ233に入った文字認識結果はテキスト取得部225及びUSBインタフェース221を介してスキャナ200に転送される。その後、当該文字認識結果がスキャナ200側の表示部213に表示され、ユーザはスキャナの手元で希望するものかどうかの確認と選択が行える。

3-3. ユーザの希望する文字認識結果は、スキャナ200より確定指示することにより、アプリケーションのカーソル位置に直接自動入力される。

#### 【0031】

（6-1-1）光学的文字認識装置について

図5を参照しながら、本発明によるイメージスキャナおよびこれを用いた光学的文字認識装置を詳細に説明する。

図5には、本発明によるイメージスキャナをマウススキャナとしたときの光学的文字認識装置の実施の形態を示すシステム構成図が示されている。図5において、パーソナルコンピュータ10は、高速双方向通信バスであるユニバーサルシリアルバス（Universal Serial Bus，以下USBと称す）に対応した、例えばマイクロソフトのオペレーションシステムであるWindows（登録商標）XP等が動作する一般的なコンピュータである。

#### 【0032】

パーソナルコンピュータ10は、ワープロソフト、表計算ソフトまたはデータベースソフト等のアプリケーションソフトが動作するとともに、文字認識ソフトがアプリケーションソフトの背後（バックグラウンド）で動作する。

すなわち、本実施の形態において、パーソナルコンピュータ10で文字認識ソフトが動作している状態でも、使用者にはアプリケーションソフトしか動作していないように見え、文字認識ソフトで認識された文字コードデータは恰もキーボードから入力されたのと同様にアプリケーションソフトに入力される。

#### 【0033】

図5には示していないが、パーソナルコンピュータ10は、本体にCRTディスプレイまたはLCD（Liquid Crystal Display）等の表示部およびキーボードが接続されている。さらに、パーソナルコンピュータ10

の本体にはUSBのコネクタが搭載され、USBケーブル50を介してマウススキャナ20に接続されている。このように、パーソナルコンピュータ10とマウススキャナ20とをUSBにより接続することで、この間での高速双方向通信が可能になるとともに、マウススキャナ20に対してパーソナルコンピュータ10より電源を供給することが可能となる。

#### 【0034】

マウススキャナ20は、パーソナルコンピュータ10のポインティングデバイスとして機能するマウスと、書類を光学的にスキャンしてそのイメージデータを出力するスキャナとが一体化された装置である。マウススキャナ20は、底面に書類を光学的にスキャンするイメージセンサ回路22を備えており、書類上の所望の位置に移動することが可能である。したがって、マウススキャナ20を文字認識を行いたい書類上の“意図する領域”に移動することで、必要な個所の文字等をスキャニングすることができる。

#### 【0035】

USBインタフェース回路24は、マウススキャナ20がスキャニングした書類上の“意図する領域”のイメージデータを、USBケーブル50を介してパーソナルコンピュータ10に送信する。

パーソナルコンピュータ10のUSBドライバ12は、イメージデータを受信すると文字認識ソフトで文字認識を行い、その文字認識結果である文字コードをUSBケーブル50を介してマウススキャナ20に送信する。

マウススキャナ20のUSBインタフェース回路24は、文字コードを受信すると、該当する文字をLCD表示部26に表示する。

#### 【0036】

このように、スキャニングした領域の文字認識結果はLCD表示部26にはほぼリアルタイムで表示されるので、認識結果の良否をスキャニングと同時に確認することが可能となる。したがって、文字認識を行う“意図する領域”の位置決め（開始箇所や終了箇所）を含め、スキャニング時点で“意図する領域”の位置の変更やイメージデータを取り込む際の設定（2値化のしきい値等）の変更を行うことができる。



## 【0037】

すなわち、本実施の形態ではパーソナルコンピュータ10の性能向上による文字認識の高速化とUSBのデータ転送速度の高速性を利用して、“意図する領域”を指定する際、表示された認識文字は、スキャナ20を僅かに動かして文字認識の開始箇所と終了箇所の位置決め等をLCD表示部26で確認できるので、“意図する領域”が取込めていなければこの時点でやり直すことが可能である。

また、パーソナルコンピュータ10が、音声合成を行うアプリケーションを搭載していれば、文字認識結果を音声出力することで使用者はLCD表示部26を確認すること無く認識結果を知ることができる。

## 【0038】

なお、“意図する領域”が取込めていることが確認でき、認識結果を確定する確定処理がマウススキャナ20で行なわれると、パーソナルコンピュータ10上で動作しているワープロソフトや表計算ソフト等のアプリケーションソフトのカーソルの位置に、この認識結果が直接入力される。

このとき、本実施の形態ではアプリケーションソフトウェア上に入力された誤認識箇所の修正をやりやすくするため、アプリケーションソフトウェアに送った文字認識結果の元のイメージデータをパーソナルコンピュータ10上に表示する。

したがって、例えば認識結果の中に“誤認識箇所”があっても、このイメージデータを参照しながらアプリケーションソフト上でキーボード等により容易にその箇所の修正を行うことが可能である。

## 【0039】

図5において、マウススキャナ20は、マウスおよびイメージセンサ回路22、USBインタフェース回路24、LCD表示部26、およびUSBマイクロコントローラ28により構成されている。なお、本実施の形態において、マウススキャナ20はスキャニングしたイメージデータを蓄積することも出来るが、通常はPC220と接続してスキャニングしたイメージデータを蓄積することなく、リアルタイムでパーソナルコンピュータ220に高速出力する。

マウスおよびイメージセンサ回路22は、パーソナルコンピュータ220のポ

インテュイティブデバイスとして機能するマウスと、書類の画像をイメージデータとして認識するイメージセンサとを含んだ回路である。マウスおよびイメージセンサ回路 22 はまた、スキャナの位置信号をマウスの位置信号として利用する位置センサ回路を備えている。さらにマウスおよびイメージセンサ回路 22 は USB インタフェース回路 24 と制御信号 (Control) の送受信を行う。

#### 【0040】

USB インタフェース回路 24 は、マウスおよびイメージセンサ回路 22 および LCD 表示部 26 と USB マイクロコントローラ 28 とのインタフェースをとる回路である。すなわち、USB インタフェース回路 24 により検出した位置情報をマウス信号としてリアルタイムに USB マイクロコントローラ 28 に通知するとともに、このコントローラ 28 よりスキャニングの開始、終了、確定等の命令を受けると、マウスおよびイメージセンサ回路 22 にこれを通知する。

また、USB インタフェース回路 24 は、USB マイクロコントローラ 28 を介して、LCD 表示部 26 に文字表示を行う制御データを受信すると、LCD 表示部 26 にこの制御データ (FFC) を出力する。この制御データにより LCD 表示部 26 には該当する文字が表示される。

#### 【0041】

USB マイクロコントローラ 28 は、マウススキャナ 20 の全体を制御する制御部と USB によりデータの送受信を行うための USB コントローラとを備えた制御回路である。USB マイクロコントローラ 28 は、USB ケーブル 50 を介してパーソナルコンピュータ 10 の USB ドライバ 12 に接続され、イメージセンサにより取り込んだイメージデータをパーソナルコンピュータ 10 に送信する。

また、USB マイクロコントローラ 28 は、パーソナルコンピュータ 10 より文字コードデータを受信すると、LCD 表示部 26 の文字表示を行うための制御データを USB インタフェース回路 24 に出力する。

#### 【0042】

マウスおよびイメージセンサ回路 22 内では、イメージセンサにより取り込んだイメージデータの一部が位置計算回路に送られ、その各々のイメージデータの

位置座標が計算される。

その後、センサー回路 22 で取込んだ複数のイメージとその各々の位置座標は、USB インタフェース回路 24 により、PC 10 に送信され、PC 10 内でソフトによって、その位置座標を基に撮影されたイメージが高速補正され、全体のイメージデータが形成される。この結果、従来問題であったハンディスキャナの蛇行や手ぶれなどによるイメージが不完全で文字が正しく認識されないという問題が解決される。

#### 【0043】

パーソナルコンピュータ 10 は、OS (Operating System) レベルで制御されるカーネル層 (Kernel Layer) の USB ドライバ 12 と、アプリケーションレベルで制御されるユーザ層 (User Layer) のユーザインタフェースとにより構成されている。

画像合成部 140 は、USB ドライバ 12 を介してマウススキャナ 20 より受信したイメージデータを取り込む。BMP フォーマット変換部 142 は、このデータをリアルタイムで例えばビットマップ (BMP) フォーマットの画像データに変換する。

#### 【0044】

文字認識処理部 144 は、当該画像データに対して日本語文字認識処理を行う。

同処理部 144 で、マウススキャナ 20 より受信したイメージデータは日本語文字の文字コードデータに逐次変換される。

一時ファイル保存部 146 は、この文字コードデータを一時ファイルに保存する。認識データ送信部 148 は、当該一時ファイルに保存された文字コードデータをほぼリアルタイムで LCD データにし、USB ドライバ 12 を介してマウススキャナ 20 に送信する。

この結果、マウススキャナ 20 でスキャンしたイメージデータは、パーソナルコンピュータ 10 によりリアルタイムでキャラクタデータに変換され、スキャンした変換結果がほとんどスキャンと同時に LCD 表示部 26 に表示される。

**【0045】****(6-1-2) マウススキャナの構造について**

図6は、マウススキャナ20の上面図および側面図である。図6に示すように、マウススキャナ20は、内部にイメージセンサ、レンズ30を備え、マウススキャナ20の全体を制御する制御回路36が配設されている。図6におけるイメージセンサ、レンズ30は、図5におけるマウスおよびイメージセンサ回路22に、図6における制御回路36は、図5におけるインタフェース回路24およびUSBマイクロコントローラ28にそれぞれ相当する。

なお、図6のイメージセンサ、レンズ30は、「光源」、「非球面レンズ」と「イメージセンサ」を一体化したイメージセンサ・モジュールである。

**【0046】**

このモジュールは操作性を考慮して、スキャン個所が常に見やすいコードの付け根に設置してある。マウススキャナ20はまた、上面にLCD表示部26が配置されるとともに、さらに2つ乃至3つのマウスボタン44が上面に配設されている。

同図の左ボタンはスキャニングをスタートするスタートボタン40として、右ボタンはLCD表示部26に表示された文字を確定する確定ボタン42として、それぞれ配設されている。

なお、特に記載はしなかったが、スキャニングの日本語／外国語／各種コードの選択、文字／画像入力 of 選択等の調整は、パーソナルコンピュータ10で行うか、マウススキャナ20にこれらの調整を行う操作ボタンを設けるか、または両方できるようにしてもよい。

**【0047】****(6-2) 実施例2**

次にイメージをとる場合について説明する。大静止画の大きさは、10mm角であるので小静止画の大きさ(2mm角)は、大静止画の20%の大きさである。画像を照合するためには、小静止画は作業中常にその総画素数の50%から30%以上の重なり合いが必要となる。この小静止画(2mm角)の重なりをとるため、大静止画は例えば1mm間隔で撮影される。

スキャン速度を 3 0 mm/秒とすると、静止画は 1 mm ( 3 3 msec ) 間隔で撮影され、各小静止画は 2 mm 角であるので小静止画の一辺 1 1 4 画素 ( 1 mm 当り 5 7 画素 ) の半分は連続した次の静止画とは約 5 0 % が常に重なり合う。またその重なり合い部分の総画素数は 6,500 画素 ( 114 × 114 ( 画素 ) × 0.5 ) となりこれが照合すべき画素サイズとなる。

#### 【 0 0 4 8 】

画素数が多い場合、パターン照合に要する計算時間が長くなり、作業に悪影響を及ぼすことが予想されるので、画像を取る場合も、文字認識のレベルまで解像度は下げて利用する。

動作手順としては、スキャン開始前に画像取り込みの設定を行うことにより、図 1 の画像利用のフローを実施する。イメージは、縦または横方向のいずれ方向に決められて配置される文字の場合と異なり、左右上下にランダムにスキャンしてイメージを取り込むことになる。

#### 【 0 0 4 9 】

##### ( 6 - 3 ) 実施例 3

実施例 1 ～ 2 ではランダムアクセス可能な CMOS センサを用いた場合を説明したが、ここではランダムアクセスが可能でないイメージセンサ CCD を用いたフリーハンド・スキャナの実施例について説明する。

本実施例のイメージセンサは、一般に市販されているインターライン型全画素読出し方式 CCD であり、メカニカルシャッター無しに、電子シャッターによる一度の露光で全画素のデータを独立に読み出せる特長を有しているとする。

#### 【 0 0 5 0 】

この CCD は正方画素 ( 縦方向と横方向の画素ピッチ一致 ) で構成されており、縦と横の解像度が一致するため、画像処理演算を容易にする。

一例として、3 3 万画素 ( 6 5 9 × 4 9 4 ) 品では、単位画素寸法が水平・垂直ともに 7 . 4 0 ミクロン、イメージ対角長 6 . 0 mm、1 / 3 型、読出し周波数 2 4 . 5 4 MHz ( 4 1 ns ) 、フレームレート毎秒 6 0 . 0 である。

このセンサの撮影サイズは 1 0 mm × 1 0 mm とし、このサイズを大静止画とし、小静止画のサイズは 2 mm × 2 mm とする。本センサのフレームレートは、

毎秒60フレームであるから、各フレームの全画素読み出しは、16.6mmかかる。

### 【0051】

図3に示す光学系で、このセンサを用いて撮影された大静止画の位置関係を計算するための小静止画を求めるが、ランダムアクセス可能なCMOSで用いたWOIはC-CDでは使えない。本センサの解像度は57画素/mm ( $=659 \times 494$  画素/mm<sup>2</sup>)であり、文字やコード認識に合わせた解像度まで、全画素の読出しの際に、画素を飛ばして小静止画の範囲の画素をメモリに収納する。この結果、撮影範囲の全画素数は $220 \times 165$ 画素となり、画像処理速度の向上を図れる。文字認識の際の動作手順を説明する。

### 【0052】

スキャン・ボタンを押し、スキャンを開始する。

1-1. まずLED光源201のパルス発光により、大静止画データ全体が撮影される。

1-2. 次に撮影された大静止画の各画素信号は、AD変換器207によりデジタルデータに変換される。まず最初に2か所の小静止画(A)(B)の画素データが読み込まれ、DSPメモリ211bへ転送される。

1-3. DSPメモリ211bへ転送されたデータはDSP211によって処理され、大静止画のA、B点の座標値への換算に用いられ、各小静止画の重心位置座標値は各フレーム撮影期間中に計算により求められる。

1-4. 次に大静止画のアドレスがスキャンされて大静止画全体が読み込まれ、画像用メモリ217へ転送される。

1-5. その後、全体の大静止画データが一定量まとめられ、位置座標(重心位置座標値、回転角度)データの添付と共に、USBを用いてPC220へ連続的に転送される。

1-6. 上記処理が繰り返され、スキャン・ボタンを離すことによりスキャンが終了する。

以下、PC220での処理は実施例2の場合と同じであるので省略する。

### 【0053】

## (6-4) 実施例 4

図 7 は本発明の実施の形態である光学端末装置の構造図である。本実施例では、携帯電話の一部としてマウススキャナを組み込み、携帯パソコンのマウススキャナとして利用したり、また、携帯電話の処理能力に余裕がある場合には、携帯電話内で文字認識処理を行ったり、さらに、図 8 に示すように携帯電話の無線公衆回線を介してサーバーと接続し、このサーバに光学端末装置にて取得したイメージデータを送信する携帯電話である。

## 【0054】

図 7 は、マウススキャナ 20 を組み込んだ携帯電話の上面図および側面図である。図 7 に示すように、マウススキャナ 20 は、内部にイメージセンサ、レンズ 30 を備え、マウススキャナ 20 の全体を制御する制御回路 36 が配設されている。図 7 におけるイメージセンサ、レンズ 30 は、図 5 おけるマウスおよびイメージセンサ回路 22 に、図 6 における制御回路 36 は、図 5 におけるインタフェース回路 24 および USB マイクロコントローラ 28 にそれぞれ相当する。

## 【0055】

なお、図 7 のイメージセンサ、レンズ 30 は、「光源」、「非球面レンズ」と「イメージセンサ」を一体化したイメージセンサ・モジュールである。このモジュールの取り付けは、意図する文字コードのスキャン個所が見やすい位置に設置してある。マウスボタンの取り付けも操作性を考慮した位置に取り付けてある。マウススキャナ 20 はまた、携帯電話に LCD 表示部 26 が配置されるとともに、マウスボタン 44 が配設されている。

なお、特に記載はしなかったが、スキャニングの日本語／外国語／各種コードの選択、文字／画像入力 of 選択等の調整は、パーソナルコンピュータ 10 で行うか、マウススキャナ 20 にこれらの調整を行う操作ボタンを設けるか、または両方できるようにしてもよい。

## 【0056】

イメージセンサ・レンズ 30 は、携帯電話のカメラと兼用することも可能であり、マウススキャナを組み込んだ携帯電話は、マウススキャナとして、USB コネクタ 50 を携帯パソコンの USB コネクタにつないで使うことができる。

さらに携帯パソコンが身近に無い場合は、取り込んだ画像情報を携帯電話の外部メモリに一時保管しておいたり、携帯電話の無線機能を用いて、取り込んだ画像情報をサーバーに送り込むことができる。また携帯電話に使われているマイクロコンピュータの能力に余裕がある場合は、パーソナルコンピュータに送り込んで処理する文字認識ソフトウェアを携帯電話に取り込んで処理することも可能である。

#### 【 0 0 5 7 】

##### ( 6 - 5 ) 実施例 5

図 8 は、本発明の実施の形態であるスキャナカメラ付携帯電話を活用することのできる各種インターネットサービスシステムの全体図である。

本実施例では、本発明の実施の形態である光学端末装置を携帯電話の一部として組み込み、携帯電話の一般公衆無線インターネット回線 8 0 0 を介し、前記光学端末装置にて取得したイメージデータを携帯電話 8 0 9 a ~ 8 0 9 c から各種サーバ 8 0 1 ~ 8 0 7 に送る。サーバ 8 0 1 ~ 8 0 7 に送られたイメージデータはサーバー内のソフトウェアにて処理される。

#### 【 0 0 5 8 】

サーバ 8 0 1 は、旅行者向け外国語文書の国際翻訳サービスを提供するサーバである。例えば、日本に来了た日本語文書が読めない外国人旅行者向けに旅行者の母国語への翻訳を行うサービスで、誰もが、いつでも、どこでも利用できる便利さが生まれる。

サーバ 8 0 3 は、カタログや新聞・雑誌に掲載された広告に記載のネットアドレスや商品コードを読み込み、通信販売・商品販売会社に直接アクセスするサービスを提供するサーバである。インフラとして普及した携帯電話が利用できる効果は大変大きい。

サーバ 8 0 5 で提供される自治体や国の各種サービスは、公報や諸通知など、さまざまな形で個人に通達されるが、電子政府・電子自治体サービスを受ける際には各サービスのアドレスを携帯電話に代表される IT 機器に入力しなければならない。このアドレス入力が容易になることは電子政府・電子自治体の利用効率に多大な効果がある。



## 【0059】

サーバ807は、情報家電へのアクセスを容易化するサーバである。情報家電の進展と共に、外出先から携帯電話809a～809cを用いてインターネットを経由し、情報家電に指示する機会が増大した。本発明によれば、携帯電話809a～809cへの情報家電アドレス入力の手力を軽減することができる。一例として、テレビ番組のGコード入力の容易化が期待できる。

以上、本発明の実施形態について示したが、本発明は前述したものだけに限定されるものではなく、その範囲を逸脱しない限り、変更及び修正されたものにも適用できることは言う迄もない。

## 【0060】

(7) 本発明の利点を以下に列挙する。

1. 発明の手ぶれ防止機能を持つフリーハンド・スキャナーによれば、書類上をスキャンする際のいわゆる手ぶれによる画像の歪みが解消され、なおかつ横方向スキャンに加えて、左右上下の全方位スキャンが可能で、完全なイメージデータを取得される。このため、使用者は手元でペン等を扱うのと同じように、手元の書類上で必要個所のみの文字認識を容易かつ確実に実行できる。特に、関心の個所がとびとびの場合、目的の個所だけの文字認識が行える。

## 【0061】

2. 文字認識を行うソフトウェアをアプリケーションソフトウェアの背後で動作させることで、認識した文字の文字コードデータをこのアプリケーションソフトウェアに直接データ入力することができる。したがって、本発明を用いれば、書類を見ながらキーボード等で文字データの入力を行うのと同じ感覚で、書類内の所望の文字を文字データ化することができるので、使用者の文字入力の作業を大幅に軽減することが可能となる。

3. 取込んだイメージデータに対する文字認識を端末あるいは携帯端末以外の装置において実行し、その認識結果をアプリケーションあるいは遠隔地にある装置で文字認識を行ってデータを装置の所望の制御部に入力することを可能とした光学的文字認識装置およびシステムが提供できる。

## 【0062】

**【発明の効果】**

以上のように、本発明の光学的文字認識装置およびシステムによれば、所望の領域をスキャンする際のいわゆる手ぶれによる画像の歪みを解消し、完全なイメージデータを取得することによって、文字認識を確実に実行できるようにするとともに、取込んだイメージデータに対する文字認識を端末あるいは携帯端末以外の装置において実行し、その認識結果をアプリケーションあるいは遠隔地にある装置で文字認識を行ってデータを装置の所望の制御部に入力することができる。

**【図面の簡単な説明】****【図 1】**

本発明による光学的文字認識装置の実施の画像処理の流れ図。

**【図 2】**

光学的文字認識装置のシステム構成図。

**【図 3】**

イメージセンサ部の撮影方法を説明するための図。

**【図 4】**

イメージセンサ部の撮影タイミング図。

**【図 5】**

本イメージセンサを用いたマウススキャナのシステム構成図。

**【図 6】**

マウススキャナの構造図。

**【図 7】**

マウススキャナを組み込んだ携帯電話の構造図。

**【図 8】**

スキャナカメラ付携帯電話が役立つ各種インターネットサービスを説明するための図。

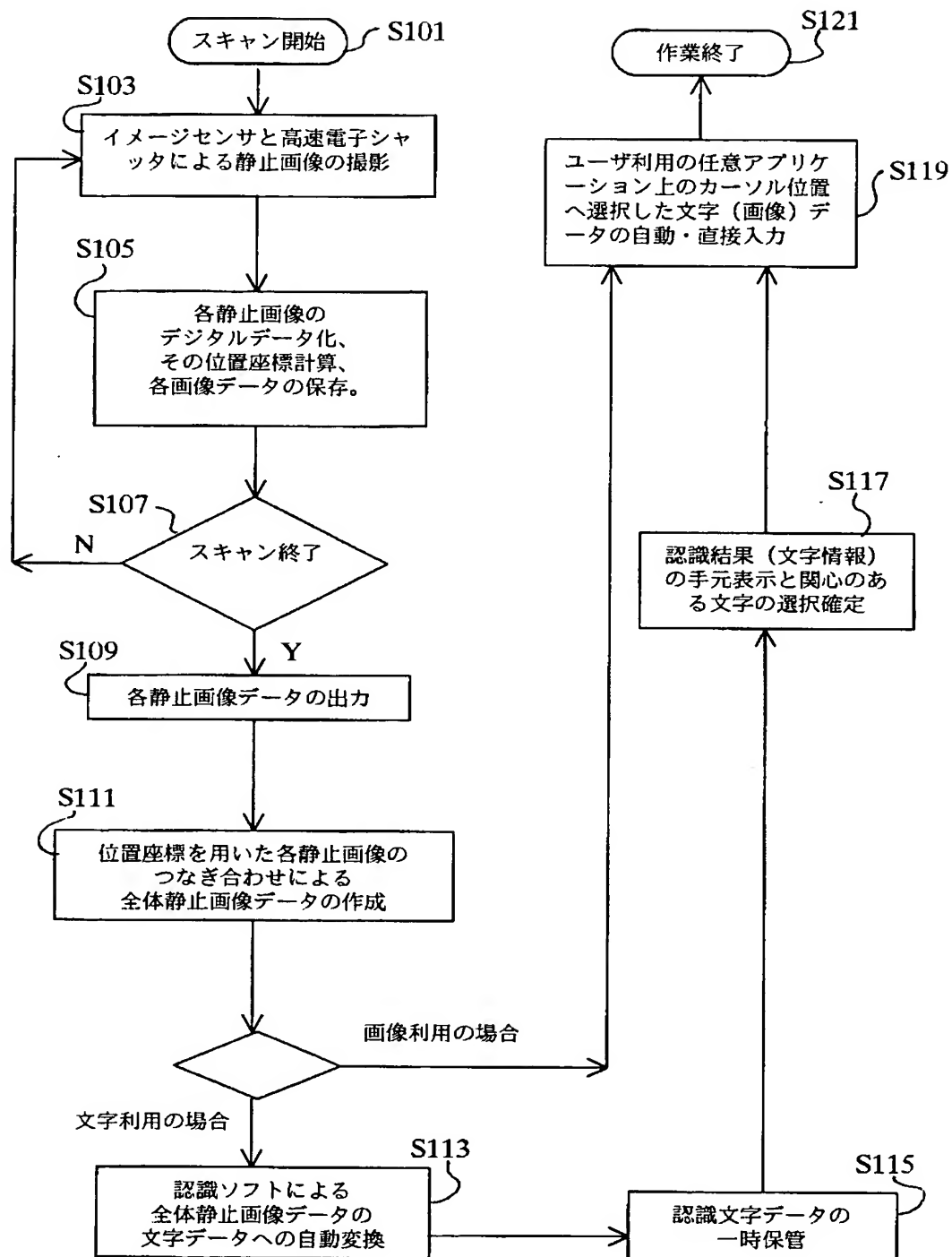
**【符号の説明】**

22	マウス及びイメージセンサ回路
24	USBインターフェース回路
26	LCD表示部

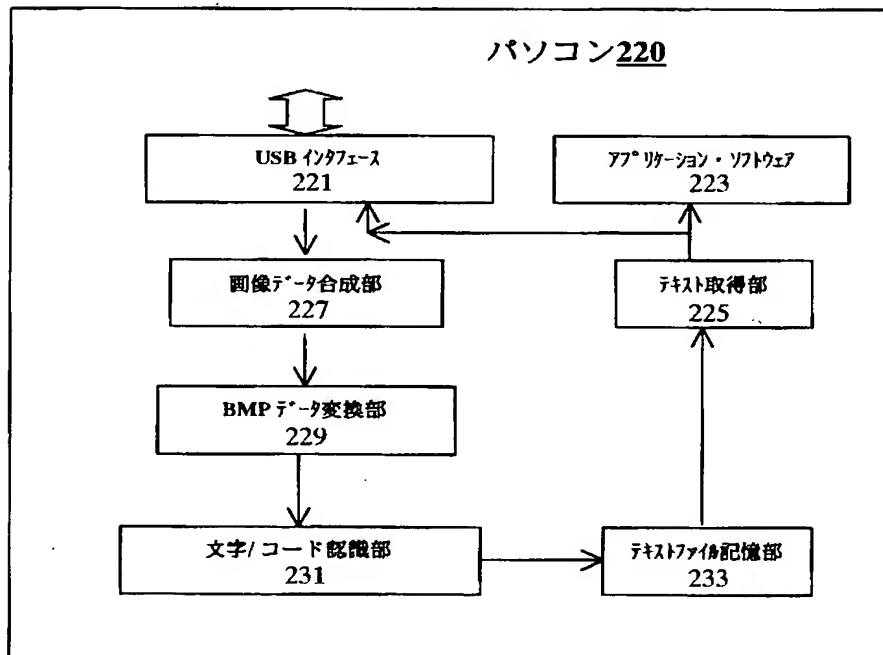
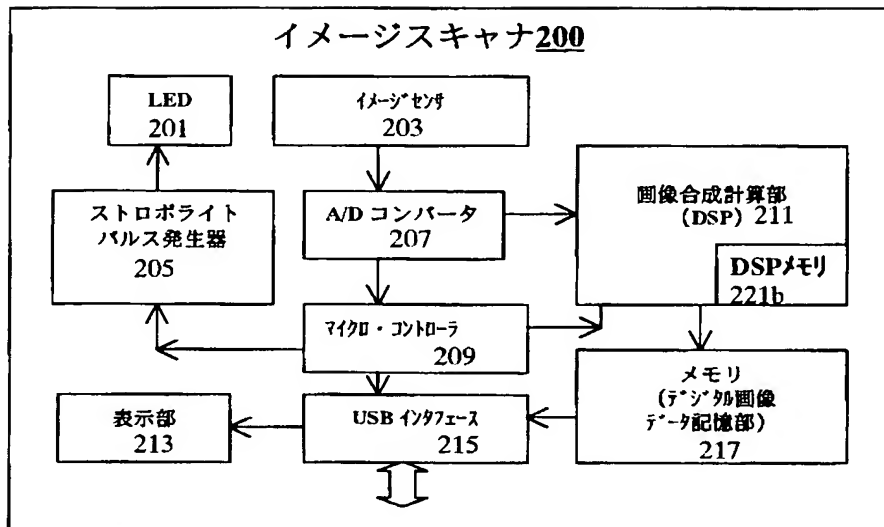
- 2 8      U S B マイクロコントローラ
- 1 4 0    イメージ合成手段
- 1 4 2    BMP フォーマット変換手段
- 1 4 4    文字認識処理手段
- 1 4 6    一時ファイル保存手段
- 1 4 8    認識データ送信手段

【書類名】 図面

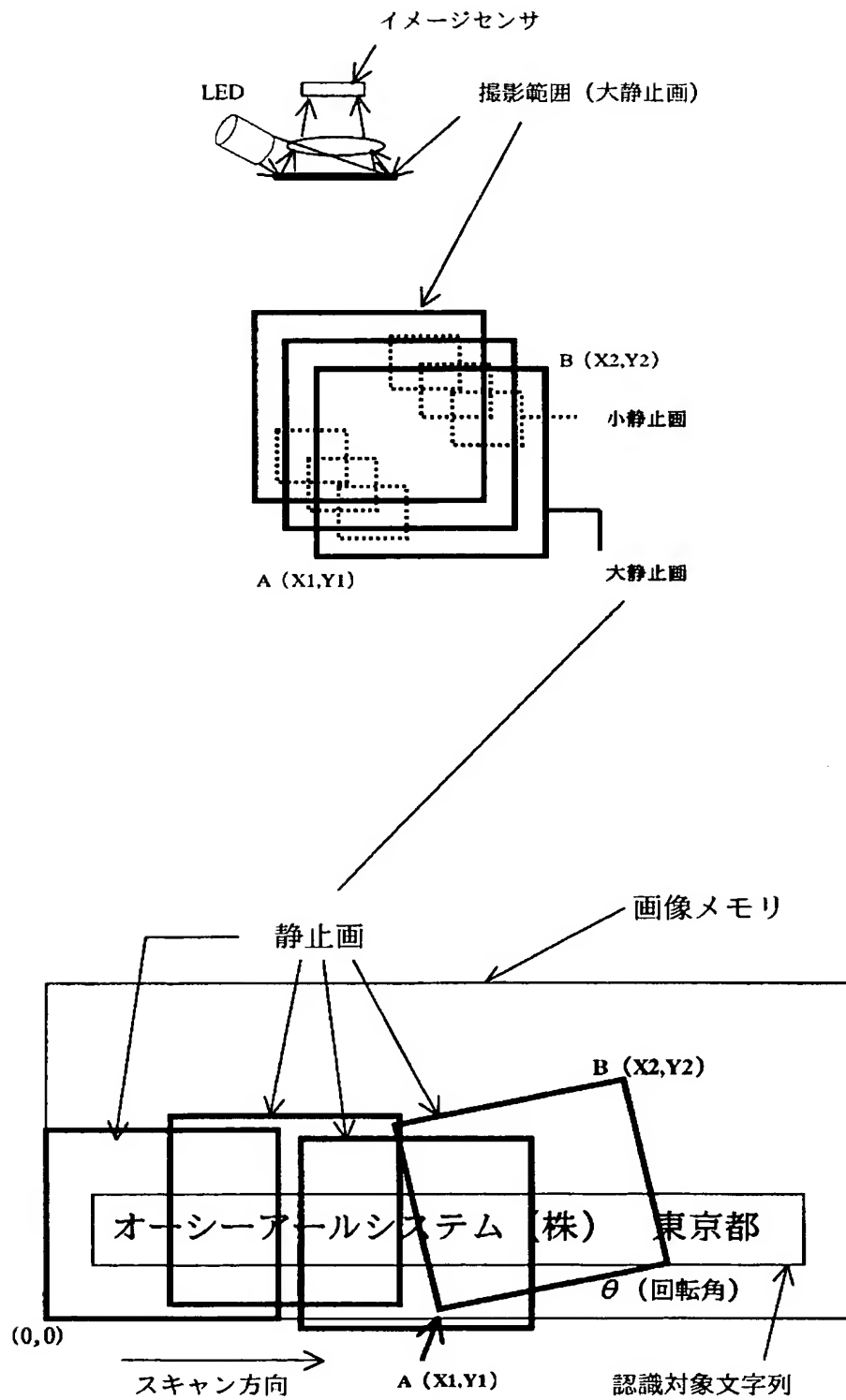
【図 1】



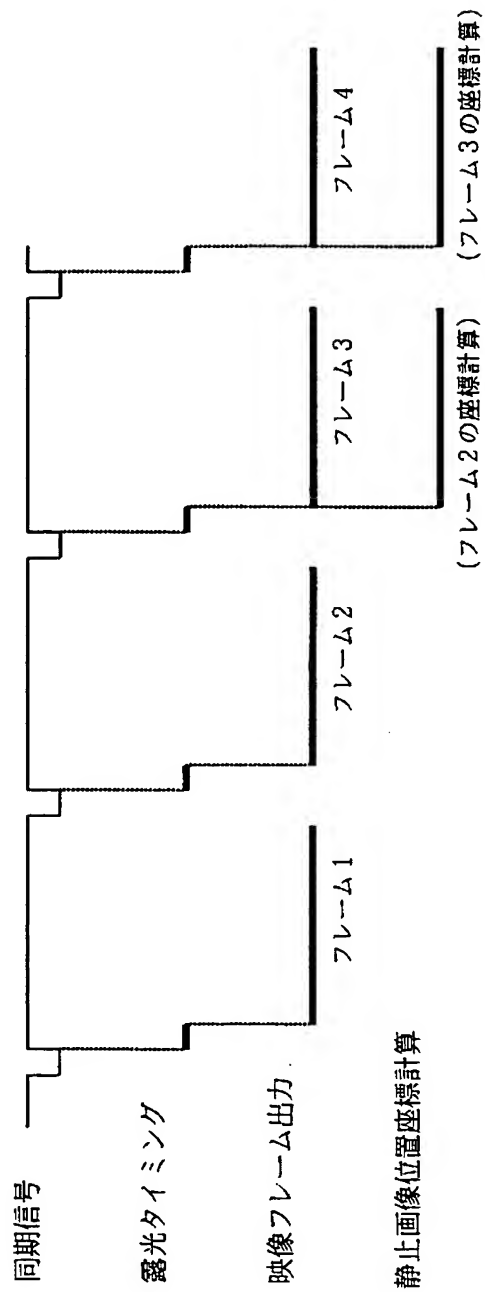
【図 2】



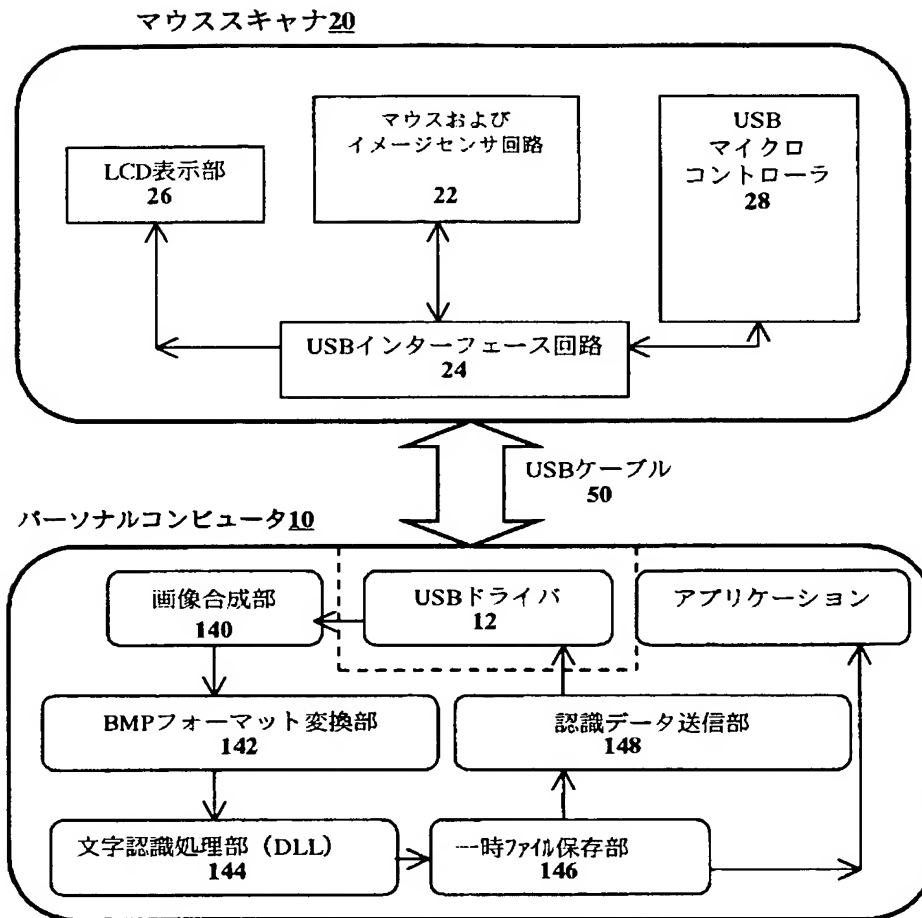
【図 3】



【図 4】

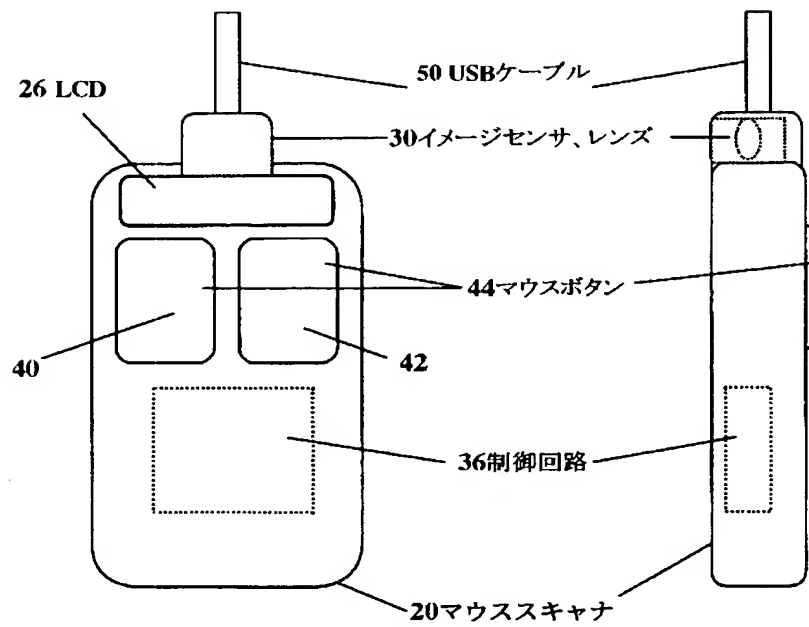


【図 5】

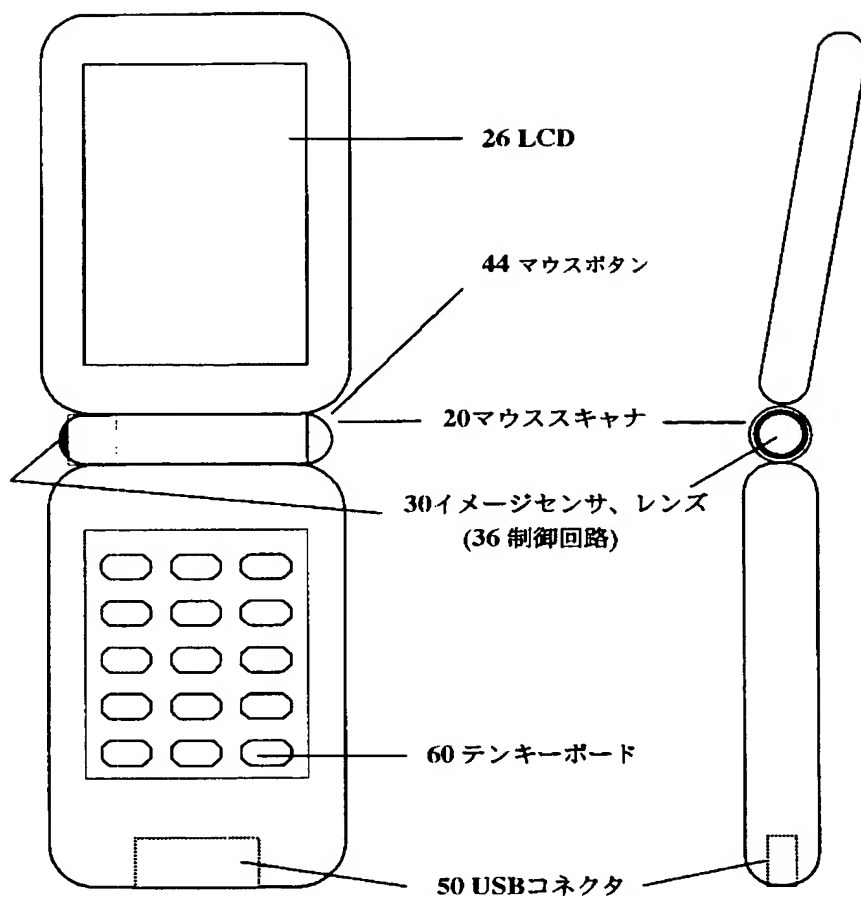




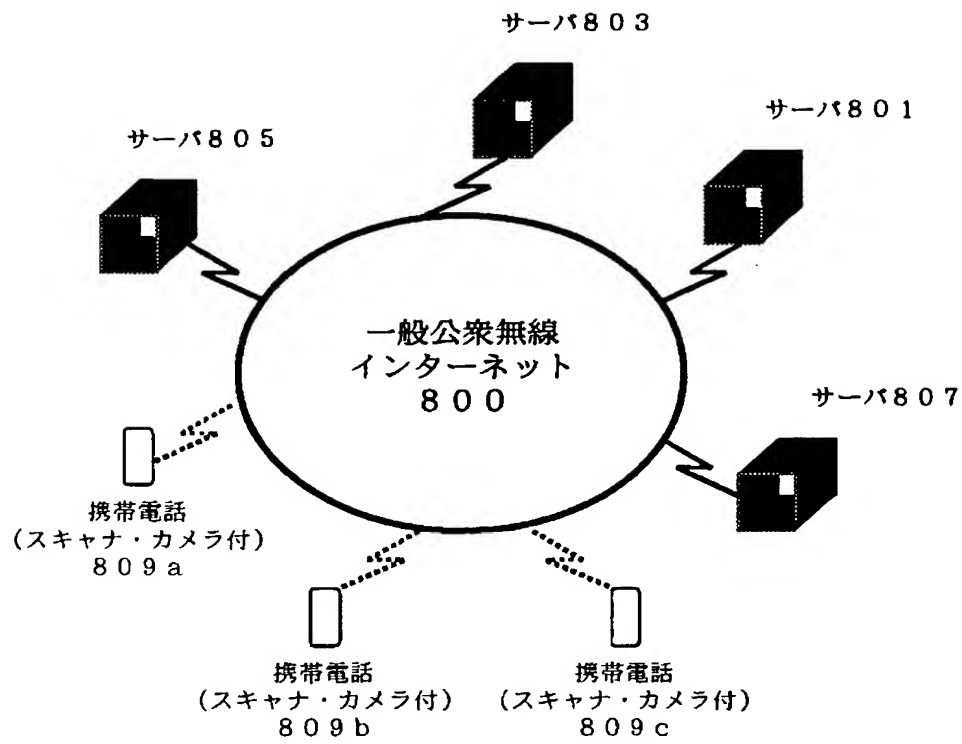
【図 6】



【図 7】



【図 8】



【書類名】 要約書

【要約】

【課題】 手ぶれによる画像の歪みを解消し、完全なイメージデータを取得することによって、文字認識を確実に実行できるようにするとともに、取込んだイメージデータに対する文字認識を端末あるいは携帯端末以外の装置において実行し、その認識結果をアプリケーション等で文字認識を行ってデータを所望の制御部に入力することを可能とした光学的文字認識装置およびシステムを提供すること。

【解決手段】 マウス機能と連動しており、書類上に書かれた関心のある個所のイメージのみを取り込め、文字やコードの認識はアプリケーションの背後で行うことができ、その認識結果はアプリケーション上の予め希望した場所へ直接入力される。スキャナ・モジュールは、マウスや携帯電話に組み込むことができ、インターネットを利用した様々な応用が考えられる。

【選択図】 図 1

特願 2 0 0 3 - 1 1 3 7 5 0

出 願 人 履 歴 情 報

識別番号

[ 5 9 8 1 4 9 3 7 8 ]

1. 変更年月日

1 9 9 8 年 1 0 月 2 9 日

[変更理由]

新規登録

住 所

東京都港区芝浦 4 - 1 3 - 3 トリニティ芝浦 1 8 0 6

氏 名

中山 光雄